REGULAR TRANSITION MATRICES

JOSEPH TOBIN

1. Introduction

In this paper, we explore properties of regular transition matrices using the guidance of Friedberg, Insel, and Spence [FIS03]. We will start off with basic definitions and then dive into theories, all of which are shown with sufficient proofs. Then, we will will deduce the relevant theorems and develop a whole swath of properties for square regular transition matrices. Finally, we will explore applications of these theorems.

We rely heavily on results shown in previous parts of section 5.3 in addition to results shown section in 5.1 and 5.2 of [FIS03]. Additionally, previous knowledge of Jordan Canonical Form is assumed.

2. Basic Definitions and Notes

Definition 2.1 (Matrix Limit)**.** Let L, A_1, A_2, ..., be $n \times p$ matrices having complex entries. The sequence $A_1, A_2, \ldots$ is said to **converge** to the $n \times p$ matrix L, called the **limit** of the sequence, if $\forall 1 \leq i \leq n$ and $\forall 1 \leq j \leq p$

$$lim_{m \to \infty} (A_m)_{ij} = L_{ij}.$$

In this case, we write

$$lim_{m \to \infty} (A_m) = L.$$

Definition 2.2 (Transition Matrix, Probability Vector)**.** We call an $n \times n$ M a **transition matrix** if it contains only nonnegative entries and all of its columns sum to 1. We call a column vector P a **probability vector** if it contains only nonnegative entries that sum to 1.

Definition 2.3 (Regular)**.** A transition matrix M is called regular if there exists an $s \in \mathbb{N}_{>0}$ such that M^s has only positive entries.

Definition 2.4 (Row Sum, Column Sum)**.** Let $A \in M_{n \times n}(C)$.

$\forall 1 \leq i, j \leq n$, define $\rho_i(A)$ to be the sum of the absolute values of the entries of row i of A or

$$\rho_i(A) = \sum_{j=1}^{n} |A_{ij}|$$

and define $\nu_j(A)$ to be the sum of the absolute values of the entries of column j of A or

$$\nu_j(A) = \sum_{j=1}^{n} |A_{ij}|$$

Then the **row sum** of A or $\rho(A)$ and the **column sum** or $\nu(A)$ are defined as:

$$\rho(A) = max\{\rho_i(A) \forall 1 \leq i \leq n\}$$

$$\nu(A) = max\{\nu_j(A) \forall 1 \leq j \leq n\}$$

Date: 27 November 2017.

3. Theorems

Theorem 1 (Theorem 5.18 pg. 298)**.** Let $A \in M_{n\lambda n}(C)$ be a matrix in which each entry is positive and let λ be an eigenvalue of A such that $|\lambda| = \rho(A)$. Then $\lambda = \rho(A)$ and $\{u\}$ is a basis for E_λ, where $u \in C^n$ is the column vector in which each coordinate equals 1.

Proof. First, note that because A has all postitive values and u has all positive values, then Au has all postive values. Therefore, because $\lambda u = Au$ has all positive values we can conclude $\lambda > 0$ and $\lambda = |\lambda| = \rho(A)$

lo we know
n eigenvalue
we don't
that it's a
tion matrix?

Now let v be an eigenvector of A corresponding to λ with coordinates $v_1, v_2, \ldots, v_n$. Then let $v_k = max(|v_1|, |v_2|, \ldots |v_n|)$ and $b = |v_k|$.

Then

$$|\lambda| b = |\lambda||v_k| = |\lambda v_k|$$

But if λ is an eigenvalue of A, then $Av = \lambda v$ and thus $\forall 1 \leq i \leq n \lambda v_i = \sum_{j=1}^n A_{ij} v_j$. Thus

$$|\lambda v_k| = |\sum_{j=1}^n A_{kj} v_j|$$

By the triangle inequality and then multiplication rules,

$$|\sum_{j=1}^n A_{kj} v_j| \leq \sum_{j=1}^n |A_{kj} v_j| = \sum_{j=1}^n |A_{kj}||v_j|$$

Since we know $b = |v_k| \geq v_i \forall 1 \leq i \leq n$ and similarly $\rho(A) \geq \rho_i(A) \forall 1 \leq i \leq n$, we know

$$\sum_{j=1}^n |A_{kj}||v_j| \leq \sum_{j=1}^n |A_{kj}| b = b \sum_{j=1}^n |A_{kj}| = b\rho_k(A) = \rho_k(A) b \leq \rho(A) b$$

But since we know $|\lambda| = \rho(A)$, we know the three inequalities above are actually equalities.

(1) $|\sum_{j=1}^n A_{kj} v_j| = \sum_{j=1}^n |A_{kj} v_j|$
(2) $\sum_{j=1}^n |A_{kj}||v_j| = \sum_{j=1}^n |A_{kj}| b$
(3) $\rho_k(A) b = \rho(A) b$

But now we can use this to show that $\{u\}$ is a basis for E_λ

□

Corollary 1.1 (Corollary 1 pg. 299). Let $A \in M_{n \times n}(C)$ be a matrix in which each entry is positiive and let λ be an eigenvalue of A such that $|\lambda| = \nu(A)$. Then $\lambda = \nu(A)$ and E_λ has dimension 1.

Proof. Consider A^T. We know by execercise 14 of 5.1 in [FIS03] that A and A^T have the same eigenvalues. Let $E_\lambda, E_\lambda{}'$ be the eigenspaces of λ corresponding to A, A^T respectively. Additionally, if $|\lambda| = \nu(A)$, then the row corresponding to $\nu(A)$ is now a column in A^T such that $|\lambda| = \rho(A^T)$. Thus A^T is a matrix in which each entry is positive with an eigenvalue $\lambda = \rho(A^T)$. Thus by Theorem 5.18 [FIS03], the basis of $E_\lambda{}' = \{u\}$, where $u \in C^n$ is the column vector in which each coordinate contains 1 and $\lambda = \rho(A^T)$. Thus $dim(E_\lambda{}') = 1$. But by exercise 13 of 5.2 of [FIS03], we know $dim(E_\lambda) = dim(E_\lambda{}') = 1$ and $\nu(A) = \rho(A^T) = \lambda$. Thus, we have shown $\lambda = \nu(A)$ and $dim(E_\lambda) = 1$.

nu, rho,
y

□

Corollary 1.2 (Corollary 2 pg. 299). Let $A \in M_{n \times n}(C)$ be a transition matrix in which each entry is positiive and let λ be an eigenvalue of A such that $\lambda \neq 1$. Then $|\lambda| < 1$ and the eigenspace corresonding to the eigenvalue 1 has dimension 1.

Proof. We know by corollary 3 of Theorem 5.16 [FIS03] that if λ is an eigenvalue of a transition matrix, then $|\lambda| \leq 1$. Thus if $|\lambda| \neq 1$, then $|\lambda| < 1$.

If A is a transition matrix, then by Theorem 5.17 [FIS03] we know that 1 is an eigenvalue.

We also know that if A is a transition matrix, then given u as a column vector in which each coordinate equals 1, then $A^T u = u$. But since u is the column vector equal to 1, each u_i in u is equal to the sum along the columns of A^T. But because $u_i = 1 \forall i$, then $\nu_i(A) = 1 \forall i$ and thus $\nu(A) = 1$.

Thus we have an all-positive-entry matrix with $1 = \lambda = \nu(A)$ and thus by Corollary 1 of Theorem 5.18 , we know $dim(E_\lambda) = 1$. fancy citing

$\square$

Theorem 2 (Theorem 5.19 pg. 298)**.** Let A be a regular transition matrix and let λ be an eigenvalue of A. Then

(1) $|\lambda| \leq 1$
(2) If $|\lambda| = 1$, then $\lambda = 1$ and $dim(E_\lambda) = 1$.

Proof. We know (1) was proved in Corollary 3 of Theorem 5.16 [FIS03].

Since A is regular, we know by definition $\exists s > 0 \in \mathbb{Z}$ such that A^s has only positive entries Because A is a transition matrix and the entries of A^s are positive, we know the entries of $A^{s+1} = A^s(A)$ are positive Suupose $|\lambda| = 1$, then we know by by Probelem 15b of section 5.1 in [FIS03] that if λ is a eigenvalue of A, then λ^s, λ^{s+1} are eigenvalues of A, A^{s+1} respectively with $|\lambda^s| = |\lambda^{s+1}| = |\lambda| = 1$.

note that we' just summing the coluymns A which have one

Because each entry of A^s, A^{s+1} is positive, we know that for any eigenvalues $\lambda*$ of A^s, A^{s+1} such that $\lambda* \neq 1$, then $|\lambda| < 1$. Thus because $|\lambda^s| = |\lambda^{s+1}| = 1$, we can conclude $\lambda^s = \lambda^{s+1} = 1$ and therefore $\lambda = 1$.

Let E_λ and $E_\lambda{}'$ be the eigenspaces of A, A^{s+1} respectively corresponding to $\lambda = 1$. Then $E_\lambda \subseteq E_\lambda{}'$ and because $dim(E_\lambda{}') = 1$, we know $dim(E_\lambda) = 1$.

whyy?

$\square$

Corollary 2.1 (Corollary Pg. 300). Let A be a regular transition matrix that is diagonalizble. Then $\lim_{m\to\infty} A^m$ exists.

Proof. We know by Theorem 5.14 [FIS03] that if $A \in M_{n\times n}(C)$ is diagonalizable and has every eigenvalue contained in S, where $S = \{\lambda \in C : |\lambda| < 1 \text{ or } |\lambda| = 1\}$, then $\lim_{m\to\infty} A^m$ exists.

Thus because A is dianolizable by assumption, we just need to show for all eigenvalues λ of A, $\lambda \in S$. But we know by Theorem 5.19 $\forall$ eigenvalues λ that $\lambda = 1$ or $|\lambda| < 1$ and thus $\lambda \in S$. Thus we can conclude that $\lim_{m\to\infty} A^m$ exists.

$\square$

Theorem 3 (Theorem 5.20)**.** Let A be an $n \times n$ regular transition matrix. Then:

(1) The multiplicity of 1 as an eigenvalue of A is 1.
(2) $\lim_{m\to\infty} A^m$ exists.
(3) $L = \lim_{m\to\infty} A^m$ is a transition matrix.
(4) $LA = AL = L$.
(5) The columns of L are identical and equal to the unique probability vector v that is equal to the eigenvalue 1.
(6) For any probability vector w, $\lim_{m\to\infty} A^m w = v$.

Proof. (1) See exercise 20 of section 7.2

(2) 1, 5.13, 5.19
(3) By theorem 5.15 [FIS03], we know L is a transition matrix if $L^t u = u$, where u is the column vector where every entry is equal to 1 Due to the equivalence, we can know L is a transition matrix is $u^t L = u^t$. But

$$u^t L = u^t \lim_{m\to\infty} A^m = \lim_{m\to\infty} u^t A^m$$

But A^m is a transition matrix which means $\forall m u^t A^m = u^t$ and thus

$$\lim_{m\to\infty} u^t A^m = \lim_{m\to\infty} u^t = u^t$$

Thus L is a transition matrix.

(4) By Theorem 5.12 [FIS03], we know $AL = \lim_{m\to\infty} AA^m = \lim_{m\to\infty} A^{m+1} = \lim_{m\to\infty} A^m A = LA$.
But $\lim_{m\to\infty} A^{m+1} = \lim_{m\to\infty} A^m = L$.
Thus $LA = AL = L$.
(5) We know $AL = L$ by (4). Let L_i be the ith column vector of L. Then because $AL = L$, we know $AL_i = L_i \forall 1 \leq i \leq n$. Thus L_i is an eigenvector of A corresponding to eigenvalue $\lambda = 1$.

Additionally, because L is a transition matrix by (3), we know $L^t u = u$ and thus $\forall 1 \leq i \leq n L_i^t u = u$. Thus $\forall 1 \leq i \leq n L_i$ is a probability vector.

But then using (1), we know that the multiplicity of 1 as an eigenvalue of A is 1. Thus all the columns of L have to be a scalar multiple of the same vector. But because every column is a probability vector, they have to be the same in order to satisify $\forall 1 \leq i \leq n L_i^t u = u$.

to prove?

Thus each column of L is equal to the the unique probability vector v corresponding to eigenvalue 1.

(6) Let w be any probability vector and $y = \lim_{m\to\infty} A^m w = Lw$. We want to show $y = v$. , If $y = Lw$, then by the corollary to Theorem 5.15 [FIS03], we know y is a probability vector.

Additionally,

$$Ay = A(Lw) = (AL)w = Lw$$

by part (4). But $Lw = y$ and thus $Ay = y$.

Therefore y is an probability vector and eigenvector of A corresponding to $\lambda = 1$ But v is the unique probability vector and eigenvector of A corresponding to $\lambda = 1$ and thus $y = v$.

Thus $\lim_{m\to\infty} A^m w = v$.

$\square$

4. Applications

References

[FIS03] S.H. Friedberg, A.J. Insel, and L.E. Spence. *Linear Algebra*. Featured Titles for Linear Algebra (Advanced) Series. Pearson Education, 2003.